The Annals of Applied Statistics
2013, Vol. 7, No. 4, 1888–1890
DOI: 10.1214/13-AOAS614E
Main article DOI: 10.1214/12-AOAS614

DISCUSSION OF "ESTIMATING THE HISTORICAL AND FUTURE PROBABILITIES OF LARGE TERRORIST EVENTS" BY AARON CLAUSET AND RYAN WOODARD

By Frederic Paik Schoenberg

University of California, Los Angeles

The similarities between terrorist strikes and earthquakes are quite striking. Clauset and Woodard's Figure 1 shows that the distribution of severity of terrorist events closely follows the Pareto or power law distribution. This distribution is also often used to describe the total scalar seismic moment, or total energy release, of earthquakes [Utsu (1999)]. It would be interesting to see if the spatial-temporal distribution of terrorist events appears to be well described by the Epidemic-Type Aftershock Sequence (ETAS) models of Ogata (1998), which are commonly used to describe earthquakes and which Mohler et al. (2011) recently used to model the spatial-temporal distribution of violent crimes in the United States.

In seismology, the *tapered* Pareto distribution has been shown to offer a somewhat better fit to some earthquake catalogs [Kagan (1993), Jackson and Kagan (1999), Vere-Jones, Robinson and Yang (2001), Kagan and Schoenberg (2001)]. Similarly, while several authors have argued that wildfire sizes follow the pure or truncated Pareto distribution [Strauss, Bednar and Mees (1989), Cumming (2001), Malamud, Millington and Perry (2005)], there is some suggestion that the tapered Pareto may fit better [Schoenberg, Peng and Woods (2003), Schoenberg and Patel (2012)]. This tapered Pareto distribution is somewhat similar to the stretched exponential considered by Clauset and Woodard, but the tapered Pareto has a pure exponential tail rather than a stretched exponential tail and therefore is typically less heavy tailed. Thus, consideration of the tapered Pareto in the context of terrorism might possibly lead to the conclusion that very large events, such as September 11, 2001, are indeed outliers.

Some might object to modeling terrorist strikes, or other events with human agency, using methods such as those employed by Clauset and Woodard.

Received July 2013.

Indeed, in addition to the usual statistical problems of misspecification, overfitting and estimation error, there is the additional issue that predictions made based on such models can be self-fulfilling, or self-unfulfilling, or have any of various intermediate feedback mechanisms based on humans' awareness of these predictions. However, Clauset and Woodard do not claim to make precise predictions of future terrorist activity based on their model, and for the purposes of describing previous terrorist strikes and forecasting broad trends of future activity, their analysis seems sensible, fair and honest.

An issue that looms large in seismology is missing data, especially with regard to historical seismicity or microseismicity [see, e.g., Kagan (2004)]. In the context of terrorism, it seems unlikely that many large events are missing from the data set. On the other hand, there must be many events that could alternately be classified as terrorist events or as conflicts, battles or even wars, depending on one's political perspective. Also, both in the case of terrorist events and earthquakes, it can be extremely difficult and subjective to determine where one event ends and another begins. Clauset and Woodard's finding that the simple Pareto distribution appears to describe the distribution of the severities of terrorist strikes so well in spite of these subjectivities is remarkable.

While it seems reasonable enough for Clauset and Woodard to treat their observations as if they were i.i.d., it should be noted that the evidence in Clauset and Woodard is far from suggesting that these events actually are i.i.d. Certainly these severities of terrorist attacks could be nearly Pareto distributed whether the events are i.i.d. or not. Indeed, the notion that these events are i.i.d. seems to contrast with the obvious nonstationarity in Clauset and Woodard's Figure 3. In seismology, earthquakes appear very obviously to arrive in clusters, both spatially and temporally. A large earthquake typically has many moderate to large aftershocks in its spatial-temporal vicinity, for example. One would anticipate terrorist events to also exhibit clustering, though perhaps not as intense as in the seismological context. It would be interesting to consider and attempt formally to test whether the sizes of terrorist events may be i.i.d. and separable, in the sense of Cressie (1991), from the locations and times of the events, as in the ETAS model; tests for this purpose were developed, for example, in Schoenberg (2004), Assuncao and Maia (2007) and Chang and Schoenberg (2011).

REFERENCES

Assunção, R. and Maia, A. (2007). A note on testing separability in spatial-temporal marked point processes. *Biometrics* **63** 290–294, 316. MR2345599

Chang, C.-H. and Schoenberg, F. P. (2011). Testing separability in marked multidimensional point processes with covariates. *Ann. Inst. Statist. Math.* **63** 1103–1122. MR2830852

Cressie, N. A. C. (1991). *Statistics for Spatial Data*. Wiley, New York. MR1127423

Cumming, S. G. (2001). A parametric model of the fire-size distribution. *Canadian Journal of Forest Research* **31** 1297–1303.

Jackson, D. D. and Kagan, Y. Y. (1999). Testable earthquake forecasts for 1999. *Seism. Res. Lett.* **70** 393–403.

Kagan, Y. Y. (1993). Statistics of characteristic earthquakes. *Bull. Seismol. Soc. Amer.* **83** 7–24.

Kagan, Y. Y. (2004). Short-term properties of earthquake catalogs and models of earthquake source. *Bull. Seismol. Soc. Amer.* **94** 1207–1228.

Kagan, Y. Y. and Schoenberg, F. (2001). Estimation of the upper cutoff parameter for the tapered Pareto distribution. *J. Appl. Probab.* **38A** 158–175. MR1915542

Malamud, B. D., Millington, J. D. A. and Perry, G. L. W. (2005). Characterizing wildfire regimes in the United States. *Proc. Natl. Acad. Sci. USA* **102** 4694–4699.

Mohler, G. O., Short, M. B., Brantingham, P. J., Schoenberg, F. P. and Tita, G. E. (2011). Self-exciting point process modeling of crime. *J. Amer. Statist. Assoc.* **106** 100–108. MR2816705

Ogata, Y. (1998). Space–time point-process models for earthquake occurrences. *Ann. Inst. Statist. Math.* **50** 379–402.

Schoenberg, F. P. (2004). Testing separability in spatial-temporal marked point processes. *Biometrics* **60** 471–481. MR2066282

Schoenberg, F. P. and Patel, R. D. (2012). Comparison of Pareto and tapered Pareto distributions for environmental phenomena. *Eur. Phys. J. B* **205** 159–166.

Schoenberg, F. P., Peng, R. and Woods, J. (2003). On the distribution of wildfire sizes. *Environmetrics* **14** 583–592.

Strauss, D., Bednar, L. and Mees, R. (1989). Do one percent of the forest fires cause ninety-nine percent of the damage? *Forest Science* **35** 319–328.

Utsu, T. (1999). Representation and analysis of the earthquake size distribution: A historical review and some new approaches. *Pure Appl. Geophys.* **155** 509–535.

Vere-Jones, D., Robinson, R. and Yang, W. Z. (2001). Remarks on the accelerated moment release model. *Geophys. J. Int.* **144** 517–531.

Department of Statistics
University of California, Los Angeles
8125 Math-Science Building
Los Angeles, California 90095-1554
USA
E-mail: frederic@stat.ucla.edu